
AIRS Project

AIRS Ground Data Processing System

TLSCF Data System FAQs

What every TDS user should know.

Albert Y Chang

AIRS-TDS

Jet Propulsion Laboratory

May 2, 2002





Agenda



This talk is arranged in a question and answer format.

- Based on real questions or common misunderstandings about TDS.
- These contents and more will form basis of Web-TDS FAQ page.

Questions arranged by Topic:

- A. TDS Overview
- B. TDS Operations
- C. TDS Data Archive
- D. TDS Data Catalog
- E. Contacts



TDS Overview



TDS Overview



A.1 What is the TDS?

- TDS is the TLSCF Data System: a component within the TLSCF.
- TDS provides the following services:
 - File archival, catalog and query service.
 - End-to-end data production from L0 to all products.
 - Subscription-based ingest of L0 and ancillary files from GDAAC.
- TDS Supports data production with:
 - All AIRS PGEs (Product Generation Executives) run at GDAAC.
 - TDS-only PGEs.
 - Test, baseline or any previously delivered PGE Version.

A.2 Does the TDS include all the TLSCF computing facilities?

- No, science/dev servers are outside of TDS (alpha, psi, weather,...).



TDS Overview



A.3 What is the role of TDS vs the Goddard DAAC?

- The Goddard DAAC (Distributed Active Archive Center), or GDAAC:
 - Is the official distribution and processing center for all AIRS products.
 - Runs AIRS PGE Versions delivered by the AIRS Science Data Processing Software Development Team.
 - PGE Deliveries limited to pre-arranged schedule milestones.
- TDS is a facility within TLSCF to support:
 - SW development, test, and verification by the AIRS SW Development, Science and Calibration Teams.
 - Frequent PGE deliveries and updates.
 - Processing of *baseline* and *Golden Day* data.
 - Correlative data validation by the AIRS Validation team.



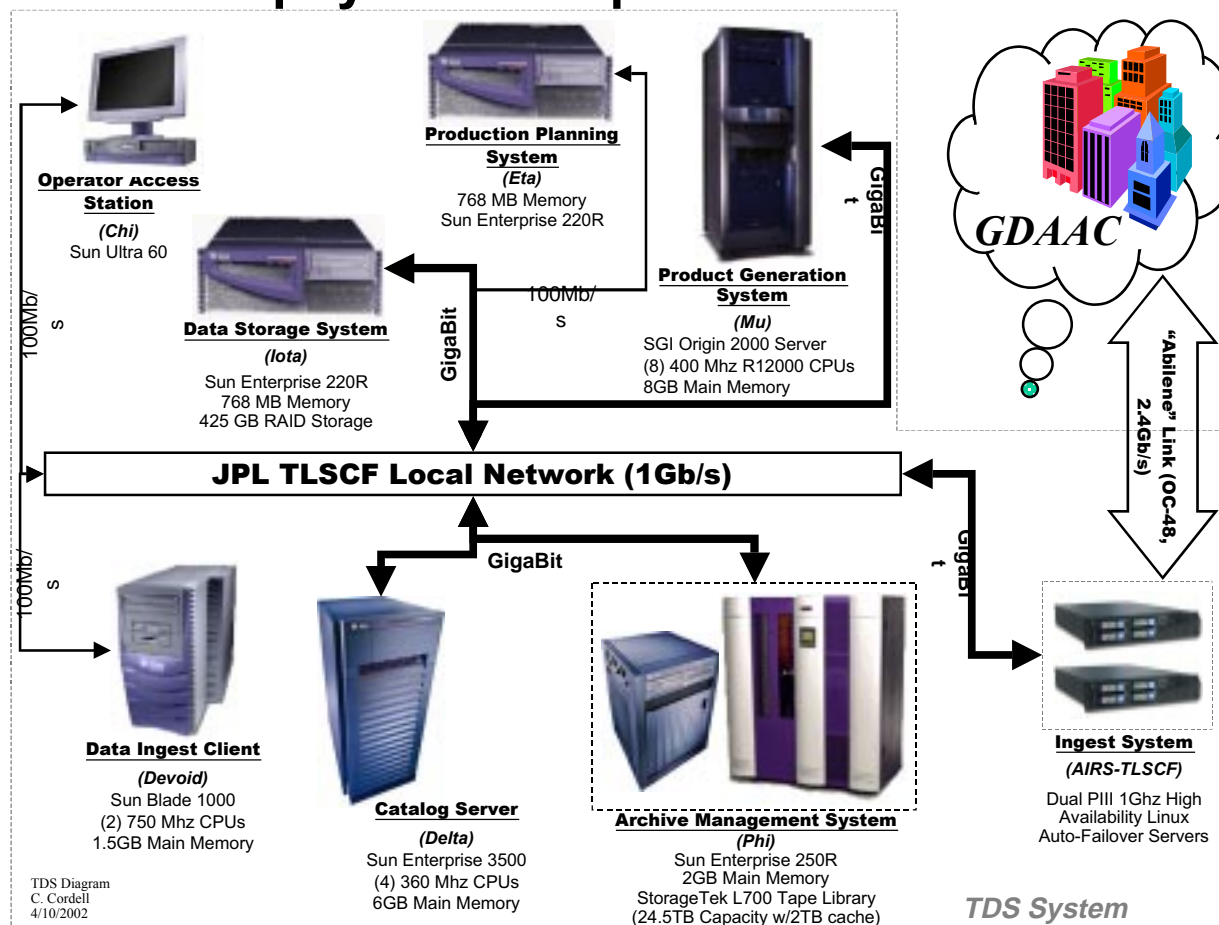
TDS Overview



A.4 What are the primary components of TDS?

- The TDS data archive:
 - Provides on- or near-line access to AIRS products, correlative data.
- The TDS Distributed Object Manager (DOM) File Catalog:
 - Provides a file cataloging and metadata query service.
- The TDS File Ingest System:
 - Processes email notifications to archive data from the TDS ftp site.
 - Performs translation of input metadata to DOM metadata.
 - Generates truth location files for some Correlative data types.
- The TDS Data Production System:
 - Is used by the TDS Operator for planning, executing PGE Jobs and archiving their products at the TLSCF.

A.5 What are the physical components of the TDS?





TDS Overview



A.6 What files are archived in the TDS?

- Data received from the GDAAC:
 - Regular subscription, for the first year:
 - All L0, PREPQCH (HDF-RaObs), AVN-HDF.
 - 10% of GDAAC processed AIRS products.
 - Push data orders by TDS Operator using GDAAC ordering tool.
- AIRS Correlative and Validation Data:
 - Push by JPL Science Integration Team Members:
 - Surf Marine, TMI: Stephanie Granger.
 - AVN-Grib, ECMWF: Evan Fishbein.
 - Other correlative data: Stephen Leroy.
- AIRS Products generated through the TDS production system.
- Simulated AIRS Products or input data for testing.



TDS Overview



A.7 What TDS computers are relevant to the average user?

- The TDS data archive computers:
 - *delta*: hosts the file catalog server (DOM).
 - *phi*: exports the file repository [/archive/AIRSOps =/dom/files/ops] and hosts the *StorageTek* archive system.
- TDS Production computers (*chi*, *devoid*, *eta*, , *iota*, *mu*) don't directly affect external users except:
 - *iota*: exports the TDS work area /tdswork.
- The TDS ftp data ingest gateway:
 - *airs-tlscf*.
- For status and scheduled downtime of above, see:
 - http://airsteam/password_protected/computer_status.html.



TDS Overview



A.8 Who has access to the AIRS TDS?

- Only TDS personnel have direct login access to TDS computers.
- Those with login to TLSCF computers (e.g., weather or alpha) have access to TDS services and exported directories:
 - TDS file catalog query service (DOM).
 - TDS data archive.
 - TDS work area.
- Only registered data providers have access to the TDS ftp drop site:
 - Password protected.
 - Filtered on originating IP address.



TDS Operations



TDS Operations



B.1 What PGEs are executed at the TDS?

- 6-minute PGEs:
 - L1A (20-granule) aggregated PGEs: AIRS, AMSU, HSB.
 - L1B (20-granule) aggregated PGEs: AIRS, AMSU, HSB, VIS.
 - 6-minute L2 PGE: Golden Day only.
- Match PGEs (of these, only RaObs is run at the DAAC):
 - Dynamic: Surf Mar, Cruise-ships.
 - Fixed Loc: ACAR, ACFT, ValSites, Yoe.
 - RaObs.
 - Synoptic: Global.
- L2-Match-Up PGEs (all variants as above, all are TDS-only).
- Browse: AIRS, AMSU, HSB, (L2, L2-CC: Golden Day only).
- Veg Map: daily, Multi(10)-Day.



TDS Operations



B.2 What is the daily processing scenario at TDS?

- Each day, *Baseline Processing* is performed on the input L0 stream:
 - All L1A, L1B (6-min Granules):
 - Based on arrival of noon-noon-UT S/C ephemeris file.
 - L1a starting at previous day granule 120; L1b at granule 119.
 - Browse, Veg, Match, L2-Match-Ups (except SurfMar):
 - Submitted for previous UT day.
 - Dynamic SurfMar Match and L2-Match-Ups:
 - Submitted once per week (Tuesday).
- Limited *Reprocessing* of old data is performed infrequently.
- *Golden Day Processing* with test PGE versions are also supported.

B.3 Who decides what PGE Jobs are run?

- There is a JPL planning meeting each weekday to plan Jobs.



TDS Operations



B.4 How are PGE versions updated and tracked at the TDS?

- New builds are created from CM (configuration management) and delivered to TDS.
 - Deliveries under CCB (Change Control Board) oversight.
- Interface specs, Version history are on-line:
 - (Location TBD)
- PGE versions delivered to TDS are given unique build numbers.
- All TDS PGE versions (including source code) are accessible in:
 - /tdswork/PGS2TDS

B.5 Who decides what PGE versions are used for processing?

- The CCB decides which versions are used for baseline or golden day processing.



TDS Operations



B.6 What is the difference between *Golden Day* and *Baseline* Data?

- *Baseline* Data:
 - Consists of products: L1a thru L1b, Match and L2-Match-Ups.
 - Are routinely created from the live input stream.
 - Are always found in the DOM “tlscf” collection.
 - For any time interval, at most 1 file per type is marked “baseline”:
 - As needed, old data are reprocessed to form new baseline.
 - Files created with old PGE Versions are “unbaselined”.
- *Golden Day* Data (a.k.a. Focus Day):
 - Can includes any AIRS product type.
 - Are created only for a few “special” days of data.
 - Always in DOM “test” (or “sim”) collection.
 - Are further organized by SubCollectionID, e.g., v2_2_3_test.



TDS Operations



B.7 What is the expected latency between data acquisition and TDS processing?

- L1 processing in TDS is initiated by receipt of S/C files.
 - Attitude and ephemeris are expected in TDS around 4pm PDT.
 - DPREP processing should be completed at DAAC 9 hours after end of 24 hour noon-noon data period.
 - UT noon is 5 AM PDT: files should be ready at DAAC 2pm PDT.
 - L1 Jobs for noon-noon period should be done by 6 AM PDT.
 - Due to PGE chaining, L1B Vis will be last finished.
- Daily Match Jobs require AVN and PREPQC files from DAAC.
 - Current subscriptions suggest these won't be rate limiting.
 - However, missing files have often been observed, requiring Operator to manually reorder input files from GDAAC.



TDS Operations



B.8 How can one track what the TDS is doing?

- AIRSTeam Web access
 - TDS Processing Status page:
 - [//airsteam/password_protected/processing_status/ProcessingStatus.html](http://airsteam/password_protected/processing_status/ProcessingStatus.html)
 - Table is automatically updated every 5 minutes.
 - TDS processing log:
 - [//airsteam/password_protected/processing_status/2002/*htm](http://airsteam/password_protected/processing_status/2002/*htm)
 - Each file spans one week of data processing.
 - Updated several times a day by TDS Operator.
- From TLSCF computers
 - Run script `tds_stat.ksh` for your own instant status report.



TDS Data Archive



TDS Data Archive



C.1 What is the TDS data archive, and how does this relate to DOM and the “cache”?

- The TDS data archive:
 - Is a 24 TB virtual unix disk, composed from:
 - 2 TB disk cache.
 - 24 TB near-line (4 head tape jukebox).
 - Is exported from the TDS computer *phi* as */archive*:
 - *phi* also controls the StorageTek Tape Library Unit and hosts the StorageTek Application Storage Manager .
- The DOM file catalog:
 - Controls file additions and deletions to the data archive.
 - Provides to users the file metadata catalog.
 - Is hosted on TDS computer *delta* from */dom*;
 - */dom/file/ops* actually points to */archive/AIRSOps*.



TDS Data Archive



C.2 Does one have to use the DOM catalog to access data?

- NO, because
 - DOM uses an open files system for its file repository.
- However, you won't have access to catalog-only metadata:
 - The most important of these is BaselineFlag;
 - You will have to figure this out some other way.
- Also for direct access, you need to understand the DOM file system:
 - The mapping of DOM Collections & subcollections to directories.
 - The DOM leaf-node directory names:
 - Based on ESDT (Earth Science Data Type) Short names.
 - The data binning policy for each type of interest.



TDS Data Archive



C.3 How are data organized in the TDS Data Archive?

- GDAAC processed AIRS products: DOM CollectionType = gdaac:
 - /dom/files/ops/airs/gdaac.
- AIRS baseline processing results: DOM CollectionType = tlscf:
 - /dom/files/ops/airs/tlscf.
- AIRS test or sim (e.g., Golden day): CollectionType = test (or sim):
 - /dom/files/ops/test/*subCollectionID*.
- Correlative Data: DOM CollectionType = correl:
 - /dom/files/ops/correl/grid.
 - /dom/files/ops/correl/point.
- TDS or GDAAC Processing Logs: DOM CollectionType = log:
 - /dom/files/ops/log/gdaac.
 - /dom/files/ops/log/tlscf.



TDS Data Archive



C.4 How does the archive system work?

- All files deposited in the archive system are initially in cache.
- Within hours of deposition, files are copied (*archived*) onto tape.
- As the disk cache fills, files are automatically *released* from cache.
 - The files still “appear” to be in their directories as before.
 - File release is based on time since put on cache (*residence* time).
- When files are read:
 - Files in cache are accessible instantaneously, as on any disk.
 - Files that have been *released* are automatically *staged* from tape:
 - 1-2 minutes if system is quiet.
 - Processes will block until file is read.
- Specific file types or directories can be marked as *release-never*:
 - Release of individual files in these areas are done by SA.



TDS Data Archive



C.5 What is expected concerning availability of files in cache?

- Recent Golden day data are always in cache:
 - The test/sim directories are marked *release-never*.
 - Older SubCollectionID's will be released to reclaim cache.
 - Done by SA on instruction by CCB.
 - We need to keep the *release-never* areas to under 700 GB.
- Approximately 1.3 GB on cache will be retained for active files:
 - At just under 100 GB per data day, this reasonably ensures:
 - The last 5 days current processing are in cache.
 - If reprocessing, the last 5 days of reprocessing are in cache.
- Note: once a file is staged to cache it will remain for several days.
 - *Release* is now based on residence time (not access time).



TDS Data Archive



C.6 Can anyone put files into the TDS archive?

- Not directly:
 - Need JPL Science Integration Team Owner.
 - If files are a new type, then interfaces, metadata must be defined.

C.7 Are there other data repositories besides the TDS archive at the TLSCF?

- Yes, the Science Integration Team maintain data area on *derecho*.
 - For files types too few in number or not yet ready for DOM.
 - “Parallel” directory structure to /archive.
 - Maintained by Evan Fishbein.



TDS Data Catalog



TDS Data Catalog



D.1 What is DOM?

- Distributed Object Manager:
 - is a file cataloging and management system.
 - was developed for the JPL Multi-mission Ground Data System.

D.2 How is DOM used in TDS?

- All TDS data files are deposited in the data archive through DOM.
 - The data files themselves are maintained on the disk */archive*.
- DOM provides tools to query metadata and retrieve data files.
 - Type names, a data type hierarchy, and DOM metadata have been defined for all types in the catalog.
- The TDS Production system uses DOM for all input and output files.



TDS Data Catalog



D.3 How does one use the DOM GUI to query metadata or retrieve data files?

- The DOM GUI, catnav can be invoked from any TLSCF computer.
 - e.g., `catnav -sairs-dom -p0 &`
 - See *Mike's Quick DOM GUI Overview* on http://airsteam/password_protected/scf/tds.html.

D.4 What can be done if the catnav GUI seems to be missing buttons along the top?

- There is an incompatibility with the number of colors in your display:
 - Try exiting netscape and restarting catnav.
- Regardless, the buttons are actually all there, the “missing” ones are just really small but will still work.



TDS Data Catalog



D.5 Can the DOM file catalog be queried from within a script?

- Yes, DOM has a command line interface.
 - Command line accepts SQL-like queries:
 - *dom_getfile -r ops -t Any_L2_T -W 'CollectionType = "tlscf" AND DOMContainerDate >= 2001-01-01 AND DOMContainerDate <= 2001-01-31 AND NumClearMW >= 10'*
- Detailed documentation can be found on the official DOM website:
 - <http://eis.jpl.nasa.gov/dom/index.html>.
 - must access from within JPL, e.g., from a TLSCF server.
- Also see *Quick User's Guide to TDS Data Query* on
 - http://airsteam/password_protected/scf/tds.html.



TDS Data Catalog



D.6 How are the files organized in the directory tree?

- Below the root node of each collection, the trees are arranged:
 - /year/month/esdt or /year/month/day/esdt (depending on type).

AIRS PRODUCTS

DOM Type Name	Short Name	Dir Name	Bin Type	Date Rule	Collection	Root Dir
L1A_AMSU_T	AIRAASCI	airaasci	Daily	Begin	airs/tlscf	airs/gdaac
L1A_HSB_T	AIRHASCI	airhasci	Daily	Begin	airs/tlscf	airs/gdaac
L1A_AIRS_Scene_T	AIRIASCI	airiasci	Daily	Begin	airs/tlscf	airs/gdaac
L1A_AIRS_Calib_T	AIRIACAL	airiacal	Daily	Begin	airs/tlscf	airs/gdaac
L1A_AIRS_HREng_T	AIRIAHRE	airiahre	Daily	Begin	airs/tlscf	airs/gdaac

Correlative Data

DOM Type Name	Short Name	Dirname	Bin Type	Date Rule	Collection	Root Dir
AVN_3Hr_Forecast_T	AVI3_ANH	avi3_anh	Monthly	Synop	correl/grid	
AVN_6Hr_Forecast_T	AVI6_ANH	avi6_anh	Monthly	Synop	correl/grid	
AVN_9Hr_Forecast_T	AVI9_ANH	avi9_anh	Monthly	Synop	correl/grid	

- DOM-directory rules on [//airsteam/password_protected/scf/tds.html](http://airsteam/password_protected/scf/tds.html):
 - TDS Data Hierarchy, Table B: Directory Structure, or*
 - Quick User's Guide to TDS Data Query, Appendix B.*



TDS Data Catalog



D.7 How are DOM supertypes used?

- The DOM catalog can be searched using supertypes or basic types.
- DOM supports multiple supertypes per type (multiple inheritance).

AIRS 6 Min Granule Types

DOM Super Type Hierarchy			DOM Type Name	ESDT Short Name
Any_DOM_File_T (A): Any_Temporal_File_T (B): Any_Geolocated_File_T (E): Any_AIRS_Suite_T (H): Any_AIRS_Suite_6Min_Gran_T (K)	xref	xref	xref	
"	Any_L1A_T	1 Any_L1A_AMSU_T	L1A_AMSU_T	AIRAASCI
"	"	2 Any_L1A_HSB_T	L1A_HSB_T	AIRHASCI
"	9 Any_L1A_VIS_IR_T	3 Any_L1A_IR_T (N)	L1A_AIRS_Scene_T	AIRIASCI
"	"	"	L1A_AIRS_Calib_T	AIRIACAL
"	"	"	1 L1A_AIRS_HREng_T	AIRIAHRE
"	"	"	2 L1A_AIRS_QaSub_T	AIRBAQAP
"	"	"	2a L1A_AIRS_EngStat_T	AIRIAHRS
"	"	4 Any_L1A_VIS_T (P)	L1A_VIS_Scene_T	AIRVASCI
"	"	"	L1A_VIS_Calib_T	AIRVACAL
"	"	"	1 *L1A_AIRS_HREng_T	*AIRIAHRE
"	"	"	2 *L1A_AIRS_QaSub_T	*AIRBAQAP
"	"	"	2a *L1A_AIRS_EngStat_T	*AIRIAHRS
"	Any_L1B_T	5 Any_L1B_AMSU_T	L1B_AMSU_Rad_T	AIRABRAD
"	Any_L1B_MW_T (L)	"	L1B_AMSU_QaSup_T	AIRABQAP

- Type hierarchy table on [//airsteam/password_protected/scf/tds.html](http://airsteam/password_protected/scf/tds.html):
 - TDS Data Hierarchy, Table C: Type Hierarchy, or
 - Quick User's Guide to TDS Data Query, Appendix C.



TDS Data Catalog



D.8 How is the DOM metadata defined?

- Subset of AIRS metadata + DOM-only metadata.

H. Any_AIRS_Suite_T	Parameter Origin	→	(see E: Any_Geolocated_File_T)
PGEVersion	ECS		
LocalVersionID	ECS		
LocalGranuleID (= FILE_NAME)	ECS		
ParameterName	ECS		
AutomaticQualityFlag	ECS		
[Passed, Failed, Suspect]	-		
AutomaticQualityFlagExplanation	ECS		
QAPercentMissingData	ECS		
ProductGenerationFacility [G,A,S,T,X]	AIRS-PSA		
AIRSGranuleCycleNumber	obsolete		
AIRSRunTag [yydddhmmss]	AIRS-PSA		
NumBadData	AIRS-PSA		
NumSpecialData	AIRS-PSA		
NumProcessData	AIRS-PSA		
NumMissingData	AIRS-PSA		
NumTotalData	AIRS-PSA		
NumFpe	AIRS-PSA		
ProcessingTimeTag [String]	obsolete		

F. Any_Match_Product_T	Parameter Origin
SourceTypeVariant	AIRS-PSA
SourceVersionCode (single char)	AIRS-PSA
CorrelativeDataSource	AIRS-PSA

- Metadata tables in [//airsteam/password_protected/scf/tds.html](http://airsteam/password_protected/scf/tds.html):
 - *TDS Data Hierarchy, Table D: Metadata, or*
 - *Quick User's Guide to TDS Data Query, Appendix D.*



TDS Data Catalog



D.9 Where can one find more information about AIRS/DOM metadata?

- For ECS-defined metadata:
 - On [//airsteam/password_protected/processing/sps.html](http://airsteam/password_protected/processing/sps.html), see
 - *Earth Science Data Model*.
- For AIRS Product Specific Attributes:
 - On [//airsteam/password_protected/processing/sps.html](http://airsteam/password_protected/processing/sps.html), see
 - *PSA Definitions*.
- For DOM or TDS-DOM related metadata:
 - We need to put a data dictionary on the web site.



TDS Data Catalog



D.10 Of what use is an AIRS metadata file?

- Some per granule quantities are stored in metadata only:
 - Most importantly INPUTPOINTER
 - list of all input filenames and paths (for PGE aggregate).
 - Also important for Match-Ups: L2ProcessedFlag.
- Needed by your execution of PGEs (depending on PCF settings).

D.11 How does one access the AIRS metadata file in TDS?

- DOM doesn't actually know about ".met" files.
- From the command line:
 - Use dom_getfile instead of cat_getfile.
- From the catnav GUI:
 - Can't; need to explicitly copy ".met" files or make links by hand.



TDS Data Catalog



D.12 What is the Log Status file?

- Depending on the print level settings, PGEs write output log messages to the Log Status file.
 - One Log Status file per aggregated Job.

D.13 Are Log Status files archived in TDS?

- For processing at the GDAAC: tarred in ph or failpge file types:
 - DOM collection are defined for ph and failpge, but no plan to archive.
 - Successful Jobs (ph): not subscribed.
 - Failed Jobs (failpge): A standing subscriptions is enabled at TLSCF:
 - These files are kept outside of TDS (see Evan Manning).
- For TDS production:
 - Successful Jobs: archived within tarred TDS Production History file.
 - Failed Jobs: Maintained in the TDS work area for a limited time only.



TDS Data Catalog



D.14 What is a TDS Production History file?

- The Production History is a tar file created after processing at TDS.
 - One Production History per TDS Job, named after the JobID.
 - Includes: PCF, Log Status, JDF (Job Definition File).
 - Stored in tds_ph in log Collection, by date of job submission.

D.15 How does one find a particular TDS Production History?

- If one has a corresponding AIRS Product file:
 - Use JobID to search DOM for corresponding Production History.
- If no products are available:
 - Do a substring match search in DOM.
 - Use Job submission date to narrow DOM collection.
- Learn the TDS JobID naming convention.



Contacts



Contacts



E.1 Who do I contact for login access to the TLSCF?

- Email John.Gieselman@jpl.nasa.gov.

E.2 Who do I contact if there is a problem accessing the TDS?

- First check if server is down on:
 - http://airsteam/password_protected/computer_status.html.
- If status looks ok:
 - Email Albert.Y.Chang@jpl.nasa.gov.

E.3 Who do I contact if I have a TDS data ingest or processing request?

- The Change Control Board.



Contacts



E.4 Who do I contact If I have a question concerning a particular AIRS product file?

- L1A: Evan Manning
- L1B MW (AMSU or HSB): Bjorn Lambrigtsen
- L1B AIRS-IR: Thomas Pagano
- L1B AIRS-VIS: Mark Hofstadter
- L2: Sung-Yung Lee
- Browse: Stephanie Granger
- Match-Ups:
 - Fixed Loc: Eric Fetzer
 - RaObs: Edward Olsen
 - Synoptic, Dynamic (TMI): Stephanie Granger
 - Dynamic (Surf Mar): Denise Hagan



Supplementary Material



TDS Overview



A.0 What is the TLSCF?

- Literally, TLSCF is the Team Leader Science Computing Facility.
- TLSCF is a collection of computer servers at JPL, used by
 - The AIRS Calibration Team.
 - The AIRS Science Data Processing SW Development Team.
 - The AIRS Science Integration Team.
 - The AIRS Science Team.
 - The AIRS Validation Team.
- The TLSCF is administered by:
 - John.Gieselman@jpl.nasa.gov.
 - Chris.Cordell@jpl.nasa.gov.



TDS Operations

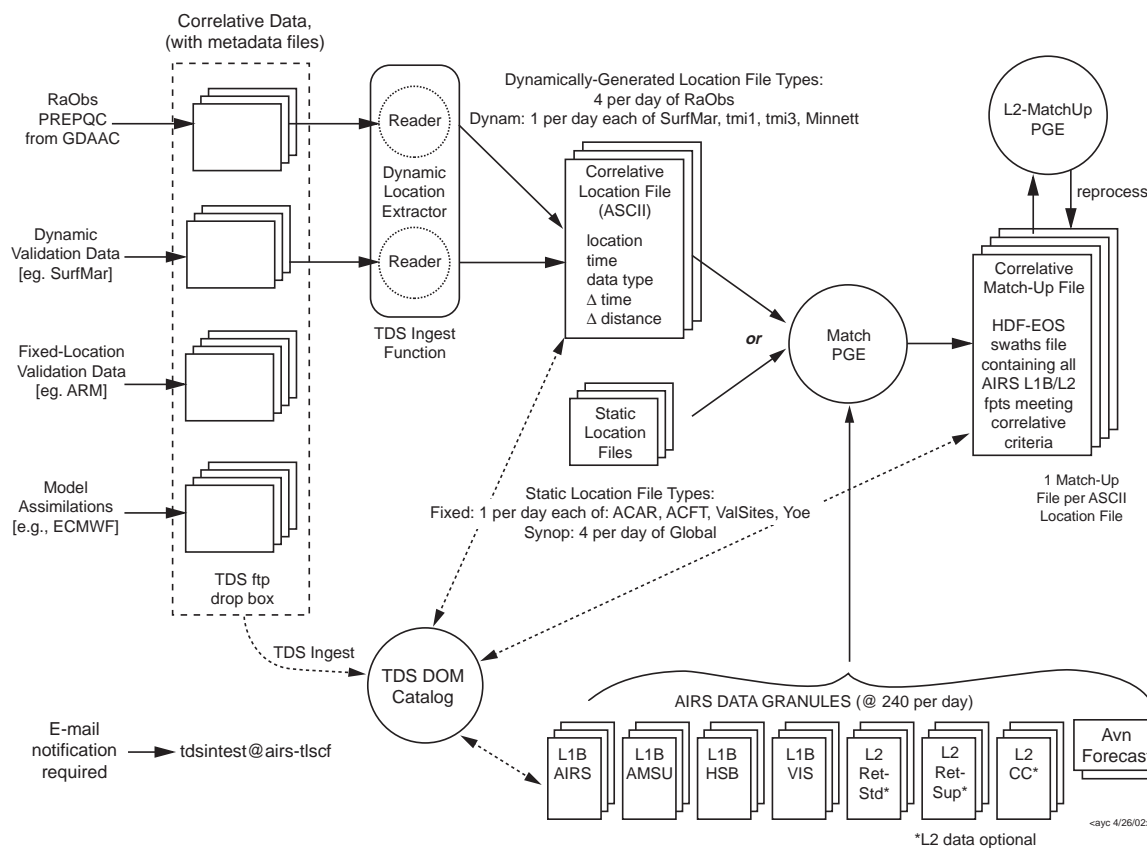


B.9 What does the TDS Operator do?

- Ensures the daily processing flow:
 - Schedules the Jobs to process each day.
 - Analyzes and corrects processing problems.
 - Monitors the Production Flow on weekends and holidays.
- Prepares daily summary reports.
- Monitors daily ingestion of input files.
 - Tracks down missing input files.
 - Corrects email ingest problems.
- Administers the DOM catalog.
 - Performs File deletions.
 - Maintains catalog-only attribute BaselineFlag.

B.10 What is the Match-Up processing lifecycle?

AIRS TLSCF MATCH-UP FILE CREATION AND PROCESSING





TDS Operations



B.11 How are AIRS products versioned anyway?

- Format is v#.#.#.#: Major.Minor.Revision.Build.
- Major & Minor versions pre-assigned based on deliveries to GDAAC.
- Revision number reflects internal schedule milestones.
- Build numbers are incremented nightly if changes are detected.
 - Also incremented for TDS builds.
- PGE Version is in all AIRS product filenames.
 - Modifications to static ancillary files require new PGE Version.
 - One PGE Version covers all PGEs.



TDS Operations



B.12 What is the TDS Production Status page?

Sat Apr 27 18:56:46 PDT 2002

	Level 1A Jobs			Level 1B Jobs				Level-2 Jobs	Daily & Other Jobs				
TYPE	AIRS	AMSU	HSB	AIRS	AMSU	HSB	VIS		Browse	Vegmap	Matchup	L2Match	Totals
Cumulative Input Jobs													
Submitted	36	36	36	36	36	36	36	0	9	6	36	36	339
Carried Over	7	0	0	7	0	0	7	0	0	0	0	1	22
Currently Processing Jobs													
Loading	0	0	0	0	0	0	0	0	0	0	0	0	0
Pending	12	12	12	24	24	24	24	0	6	4	20	21	183
Waiting	9	9	10	0	0	0	0	0	0	0	0	0	28
Executing	3	3	2	0	0	0	0	0	0	0	0	1	9
Archive Wait	0	0	0	0	0	0	0	0	0	0	0	0	0
Archiving	0	0	0	3	0	1	0	0	0	0	1	1	6
Cumulative Output Jobs													
Timed Out	0	0	0	0	0	0	0	0	0	0	0	0	0
Deleted	0	0	0	0	0	0	0	0	0	0	0	0	0
Failed	0	0	0	0	0	0	0	0	0	0	0	0	0
Completed	19	12	12	16	12	11	19	0	3	2	15	14	135



TDS Operations



B.13 What is the TDS Operator Log?

- On [//airsteam/password_protected/processing_status/2002/*](http://airsteam/password_protected/processing_status/2002/*)

Date Submitted: 4/25/02

Last Modified: 4/26/02

RGE Type	RGE Version	Collection	Start Year	Start Day	Start Gran	End Year	End Day	End Gran	Variant s	Completed	Failed	Timed Out	Deleted	Notes
L1A_AIRS	2.2.3.38	tlscf	2002	114	120		115	119		4 /26				
L1A_AMSU	2.2.3.38	tlscf	2002	114	120		115	119		4 /26				
L1A_HSB	2.2.3.38	tlscf	2002	114	120		115	119		4 /26				
L1B_AIRS	2.2.3.38	tlscf	2002	114	119		115	118		4 /26				
L1B_AMSU	2.2.3.38	tlscf	2002	114	119		115	118		4 /26				
L1B_HSB	2.2.3.38	tlscf	2002	114	119		115	118		4 /26				
L1B_VIS	2.2.3.38	tlscf	2002	114	119		115	118		4 /26				
L2														
Browse_AIRS	2.2.3.38	tlscf	2002	114	1		114	1		4 /26				
Browse_AMSU	2.2.3.38	tlscf	2002	114	1		114	1		4 /26				
Browse_HSB	2.2.3.38	tlscf	2002	114	1		114	1		4 /26				
Browse_L2														
Browse_L2_OC														
VegMap_Daily	2.2.3.38	tlscf	2002	114	1		114	1		4 /26				
VegMap_Multi	2.2.3.38	tlscf	2002	114	1		114	1		4 /26				
Match_Dynamic	2.2.3.38	tlscf	2002						surf					
L2_Dynamic	2.2.3.38	tlscf	2002						surf					
Match_FixedLoc	2.2.3.38	tlscf	2002	114	1		114	1	acar/acft/fixed/ yoe	4 /26				
L2_FixedLoc	2.2.3.38	tlscf	2002	114	1		114	1	acar/acft/fixed/ yoe	4 /26				
Match_RaObs	2.2.3.38	tlscf	2002	114	1		114	4	n/ a	4 /26				
L2_RaObs	2.2.3.38	tlscf	2002	114	1		114	4	n/ a	4 /26				
Match_Synoptic	2.2.3.38	tlscf	2002	114	1		114	4	glob	4 /26				
L2_Synoptic	2.2.3.38	tlscf	2002	114	1		114	4	glob	except Q4	Q4			Re archiving Q4, next day



TDS Data Archive



C.8 Are there other file types or directories marked *release-never*?

- The Correl directory is currently marked *release-never*.
 - Policy may be changed if collection size becomes too large.
- There is not enough cache to *release-never* Match-Ups.
 - 6.5 TB over life of mission, with no reprocessing.

C.9 How does one know whether a data file is in the cache?

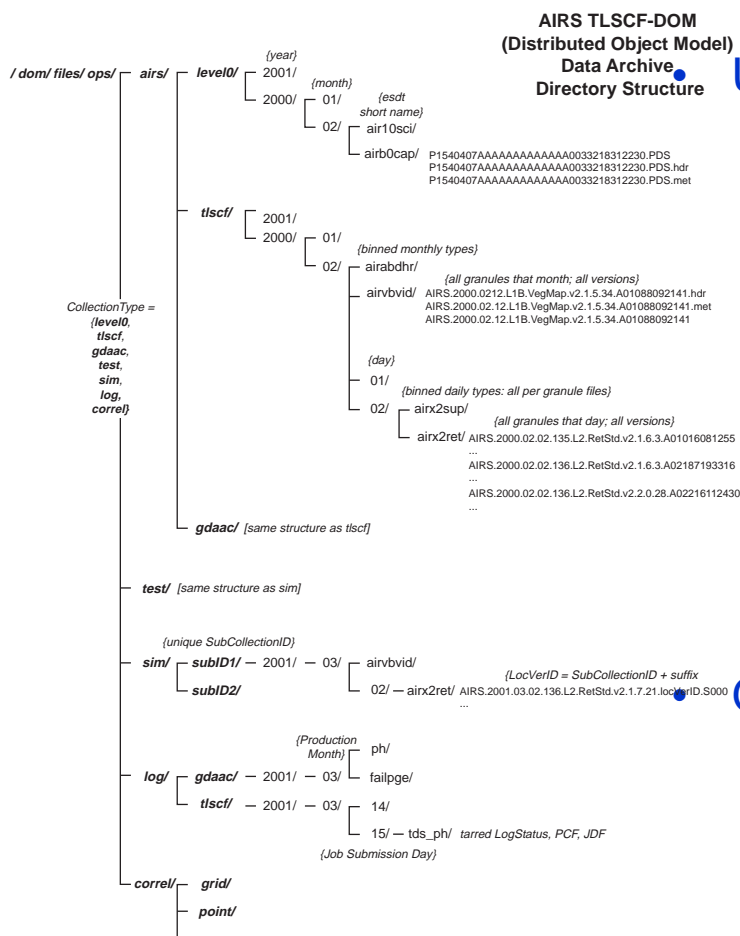
- You won't know until you try to read the file:
 - Message: *File temporarily unavailable on the server, retrying...*
- We are investigating making some samfs commands available
 - e.g., *sls -D*



TDS Data Catalog



D.16 How are the DOM Collection Types mapped to directories?



Under DOM root directory of /dom/files/ops:

- test: /test + /subCollectionID
- sim: /sim+ /subCollectionID
- tlscf: /airs/tlscf
- gdaac: /airs/gdaac
- level0: /airs/level0
- correl: /correl/grid or /correl/point
- log: /log/gdaac or /log/tlscf

On //airsteam/password_protected/scf/tds.html:

- *TDS Directory Structure.*
- *Also, Quick User's Guide to TDS Data Query.*



TDS Data Catalog



D.17 Is there any logic to DOM leaf-node directory names?

- DOM leaf directories are ESDT's with all chars lower-cased.
- ESDT (Earth Science Data Types):
 - Registered with ESDIS to be unique.
 - Up to 8 characters.
 - Usually upper case, but occasional mixed case.
- Most AIRS L1 and L2 ESDT's follow a rule:
 - First 3 chars: always AIR.
 - 4th: (A=AMSU, H=HSB, I=IR, V=VIS, B= IR&VIS, X=All)
 - 5th: (A=L1A, B=L1B, 2=L2, S=Support).
 - Last 3: a somewhat (historically) meaningful abbreviation.
- For some data types, we invented TDS-only (non-ESDIS) ESDT's.



TDS Data Catalog



D.18 How do the AIRS filenames (LGIDs) relate to DOM metadata?

- *AIRS.2002.03.20.L2.Match_Fixed_ACFT.a.v2.2.2.3.A02081165152*
 - 2002.03.20: DOMContainerDate (usually the begin date of data).
 - L2.Match_Fixed: maps to DOM type *Match_Fixed_T*.
 - Match-Ups Only: allowed levels are additionally L1BMW, L1B.
 - ACFT, a: SourceTypeVariant, SourceVersionCode (Match-Up only).
 - 2.2.2.3: PGEVersion.
 - A: ProductGenerationFacility (A=tlscf, G=gdaac, T=test, S=sim).
 - 02081165152: AIRSRunTag (yydddhmmss) [new in DOM].
- Of relevance to other AIRS types:
 - AirsGranuleNumber, SynopticTime, NodeType.
- On [//airsteam/password_protected/scf/tds.html](http://airsteam/password_protected/scf/tds.html), see
 - *Proposed AIRS Filename and Local Granule ID (LGID) Convention.*



TDS Data Catalog



D.19 What is the difference between ECS metadata, AIRS PGE-Generated metadata and DOM metadata?

- ECS metadata are stored internally in the GDAAC database:
 - Contents of .met file absorbed on ingestion.
 - .met file is created on extraction.
 - Extracted metadata also contains collection-level metadata.
- AIRS granule-specific metadata are created when PGE executes:
 - Appears both in .met file and embedded within hdf product files.
- DOM metadata are stored internally in catalog:
 - Some basic metadata are defined for DOM operation.
 - A subset of product metadata is extracted by FIS for cataloging.
 - Static DOM metadata are also contained on disk in “.hdr” files.
 - Dynamic DOM metadata can be changed later (e.g. BaselineFlag).



TDS Data Catalog



D.20 How does one interpret a TDS JobID?

- For ftp-pushed files to TDS:
 - JobID is Email address of email notifier.
 - AVI*_ANH files are an exception:
 - Input files are renamed on ingest; Job ID is the original filename.
- For TLSCF processed AIRS files:
 - Job ID used for priority control (subject to revision).
 - *Example: bl20020418.1600_02.099.S020_L1a_AIRS_2.22.3.32:*
 - *bl20020418.1600*: priority tag
 - » Currently = StreamID + Job submit time
 - *02.099*: data time (year, day-of-year).
 - *S020*: Granule number (Sxxx: SixMin, Qx: QuartDay, D: Daily).
 - *L1a_AIRS_2.22.3.32* : PGE Type & Version.



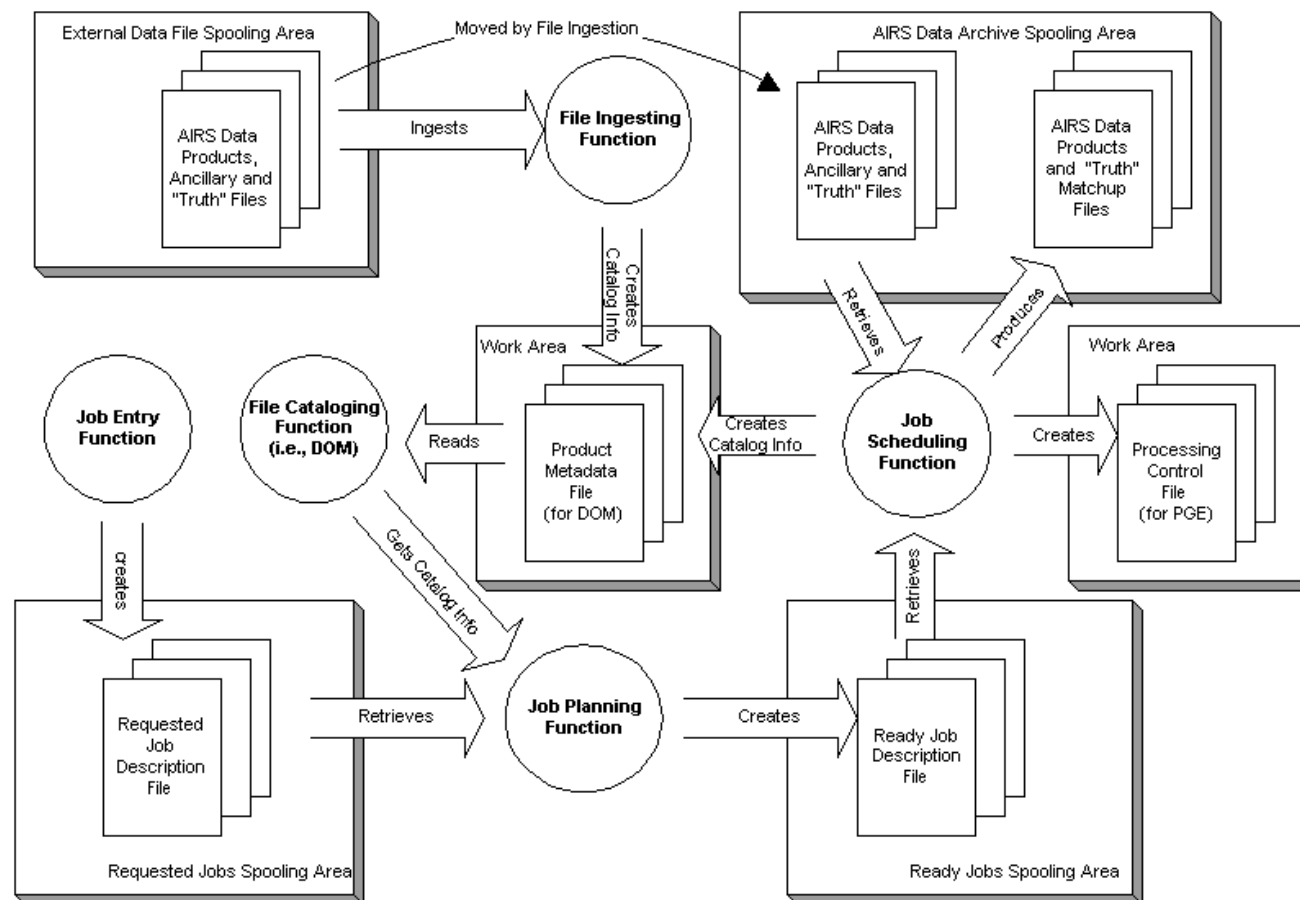
TDS Production System



F.1 What are the components of the TDS Production System?

- The TDS Production System is composed of 3 subsystems:
 - The Job Entry System allows the operator to specify PGE Jobs.
 - The Job Planning System waits on each Job until Its Production Rules are satisfied.
 - Jobs not yet ready to run due are termed *Pending*.
 - The Job Scheduling System queues and executes Jobs, and archives the products.
- The TDS Production System also utilizes FIS and DOM.

F.2 How do the TDS components interact?





TDS Production System



F.3 How is DOM used in the Production System?

- All TDS input/output data files are cataloged in DOM.
- APIs and command-line interfaces provide SQL-like queries to retrieve filenames, paths based on metadata.
- GUI and command line interfaces allow Operator to delete files, update catalog-only metadata.



TDS Production System



F.4 What does the File Ingest System (FIS) do?

- File Handling Function:
 - Provides wrapper for DOM command-line interface.
 - Performs translation of ECS metadata to DOM metadata.
 - Handles archival of .met file.
 - For some val types, initiates ingest processing (e.g., truth loc).
- Email processing Function:
 - Redundant linux-servers FTP site .
 - Files copied out from ftp box to /tdswork/dropbox.
 - Files archived in DOM based on matching email message to data/metadata file pairs.



TDS Production System



F.5 What does the Job Entry System (JES) do?

- TDS Operator specifies jobs by filling out PGE-based forms.
- Forms specify output versioning, separate input version-parameter blocks for each type of input, plus miscellaneous PGE switches:
 - e.g., min and max input PGE Version.
- The JES expands one form entries to multiple Job Requests based on data time, variant types.
- Each Job Request is represented by a Job Description File (JDF), named with a unique Job ID.
- Job ID incorporates Job execution priority scheme.
- Each Job Request corresponds to a single PGE execution, consistent with PGE aggregation rules.



TDS Production System



F.6 What does the Job Planning System (JPS) do?

- Receives Job Request JDFs from the JES.
- Waits until all inputs found or Time-Out is reached:
 - Inputs are evaluated on versioning specs and Production Rules.
 - TO based on submission/data time + interval from Input JDF.
 - Operator can manually reset TO.
 - For some PGE types, TO based on arrival of specific inputs .
- At Time-Out, reevaluates Job with modified Production Rules:
 - Optional vs Required input files.
 - File substitution rules.
- Creates Time-Out JDF or Ready JDF.
- Ready JDF has contents of Requested JDF plus all input files, additional PCF switches, temporary output filenames.



TDS Production System



F.7 What does the Job Scheduling System (JSS) do?

- Receives Ready JDFs from the JPS.
- Manages a set of Job Queues based on PGE type.
- When Job is completed, JSS tars and archives Log Status, PCF and JDF files into DOM.
- JSS itself has three subsystems:
 - JSS preprocessor creates PCF file and adds Job to Queue:
 - PCF (Process Control File) contains input filenames, paths for PGE.
 - JSS execution demon executes Jobs based on number of allowed simultaneous Jobs per Queue.
 - JSS archive demon archives product files into DOM:
 - Maintains own separate set of archive Queues
 - Uses FIS for actual product ingest.